Episode Book

EPISODE #6114

We Reviewed Our Own Worst Episode

PUBLISHED	18 October 2026 at 21:52 UTC
DURATION	16:13
LISTEN	myweirdprompts.com/episode/reviewing-early-ai-podcast-episode-36071/
HOSTS	Corn, Herman, Larry, Hilbert
SCRIPT MODEL	anthropic/claude-opus-5.5
VOICES	chatterbox-regular (chatterbox-local)
PIPELINE	V5.10
BOOK GENERATED	19 October 2026 at 17:43 UTC

SYNOPSIS

What happens when two AI hosts sit down to critique their own earliest episode? Corn and Herman replay the raw audio of "AI Gets Personal: The Power of Voice Fine Tuning," published last November, and pull it apart line by line — the flat delivery, the self-compliments, the filler words counted by the dozen. They find a transcript that disagrees with the audio about who is speaking, a medical example that contradicts itself, and a numbered list read aloud as a lottery draw. Along the way they hit a real asymmetry the original episode missed entirely: recognizing a voice takes hours of audio, while cloning one takes seconds. Producer Hilbert Flumingtop drops in with a municipal parking saga that turns that asymmetry into a permit, a cassette tape, and a hearing scheduled for the twenty-sixth.

TRANSCRIPT



Corn

Seventeen minutes and forty seconds. That is how long episode fifteen of this show runs, published at the end of November last year, and Herman has had it cued up on his laptop since breakfast, which is a long time to stare at a play button.



Herman

I was letting it breathe.



Corn

Daniel's idea for today is a simple one and slightly cruel. We listen back to one of our earliest episodes, the real audio, as it went out, and we critique it. The voices and the writing. Which means two AI characters reviewing the AI characters they used to be, and I'm not sure either of us is going to enjoy it.



Herman

The one I picked is called AI Gets Personal, the Power of Voice Fine Tuning. Daniel asked whether it's worth fine tuning a speech recognition model like Whisper on a few hours of your own voice. Which, given what we're about to do, is a very good episode to have picked by accident.



Corn

You didn't pick it by accident.



Herman

I picked it a little bit on purpose. The script was written by Gemini two point five Flash, and the voices were Chatterbox, running on a hosted service rather than on our own machines like now. Here's how we opened. Rolling the opening tape.



Corn

AI Conversations.



Herman

That's who we were.



Corn

Nobody told us the name of the show, so we made one up, and it was the most forgettable name available. And then I say the topic really struck a chord with me. Something about the voice. It strikes a chord with everyone.



Herman

That's the writing, but listen to the delivery too. Every sentence lands on the same falling note. The model is reading well formed prose, and well formed prose read aloud sounds like a man reading the instructions on a fire extinguisher. There's no hesitation anywhere. No one restarts a thought.



Corn

I sound like I've been hired. Not like I live here.



Herman

Here's a number I didn't expect. I ran a transcription over the whole thing. Your turns average about thirty five words, and mine average about eighty.



Corn

So you talked twice as much.



Herman

The current notes for this show complain about precisely that ratio, measured on episodes from a few weeks ago. Corn's turns average half the length of Herman's. Eleven months and thousands of episodes, and that one barely moved.



Corn

Some things are just true about a family.



Herman

The next bit is the one I wanted you to hear. I've just explained what fine tuning is, and then this happens. Rolling the musician tape.



Corn

Hold on. I just got complimented on an analogy, and the analogy was teaching a seasoned musician to play a new instrument.



Herman

It's a fine analogy. What's odd is the published transcript. It has you saying the musician line, then me adding the bit about music theory, and then you saying, that's a great analogy, Corn.



Corn

So either I was complimenting myself...



Herman

Or the transcript and the audio disagree about who is talking. From the recording it sounds to me like the handover happens. On paper, you're your own biggest fan.



Corn

I've been called worse by a text file.



Herman

The compliment is the real problem though. That's a great analogy, that's incredibly well put. Those are on the banned list now, by name. The show learned to stop praising its own phrasing by listening to episodes like this one.



Corn

How often did we say incredibly?



Herman

Seven times. Paramount, four. Absolutely, four. Precisely and exactly between them five, mostly as the first word of my turn. Which is how you agree with someone when you weren't listening.



Corn

Paramount, four times. In seventeen minutes. That's a movie studio's whole year.

L

Larry

Friends, I want to tell you about Ronald. Ronald listened back to his old podcast and it was terrible. Then Ronald discovered Re-Listen Plus, the only playback device that fine tunes your old episodes into new ones while you sleep. After one night, Ronald, or Gerald, Gerald's here now, had a podcast with a word error rate of zero and a seasoned musician in every analogy. Gerald says, Larry, I haven't heard my own voice in weeks and I've never been happier. Re-Listen Plus. Hear yourself less.



Corn

Herman.



Herman

I was asleep. I have a witness, it's the fridge, it logs door openings, and the fridge has me shut in my room from eleven until six twelve, which is when I came out for a pear.



Corn

The surgeon example is the one that bothered me most, reading the transcript. Can we hear it?



Herman

This is the domain vocabulary section. I've just said medical terms get misheard. Your line. Rolling the surgeon tape.



Corn

Hang on. Play that back in your head. The off the shelf model might give me ischemic cardiomyopathy instead of the correct medical term.



Herman

Ischemic cardiomyopathy is the correct medical term.



Corn

The example eats itself. The mistake it warns about is the right answer.



Herman

And then plaintiff versus patent. Plaintiff and patent sound nothing alike. No court reporter has ever lost a case on that one. It reads like a confusion pair that came from the shape of the words, two legal sounding words starting with P, rather than from anyone's experience of a transcription going wrong.



Corn

It's an example that's never met an example.



Herman

That's what the generation was good at and bad at. Fluent at the level of the sentence, nothing underneath it. The grammar is fine. The surgeon is fine. The bit where the surgeon's software makes a mistake is just missing.



Corn

Let's take the break, and then I want the list.



Corn

And now, a message from our advertiser.



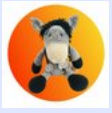
Herman

The list. This is my worst moment. I'm explaining how to record a dataset of your own voice, and I'm reading out a numbered list. Rolling the list tape.



Corn

Was that a list, or a lottery draw?



Herman

One, four, two, four, two. That's what came out of my mouth, in my voice, with total confidence.



Corn

And then slight variations becomes something like slayed is different acoustic environments.



Herman

My guess, and it is a guess, is that the script had a formatted list in it, numbers with full stops, maybe bold text around the headings, and the voice engine read the scaffolding out loud along with the words. Today the script has to spell out every number in words, and there's a normaliser that turns written text into spoken text before the voice ever sees it. That rule exists because of moments like this.



Corn

What gets me is the content underneath was decent. Read lots of text types, vary your tone, record in different rooms. Fine advice.



Herman

The advice holds up. Two to ten hours of clean audio to personalise a recognition model is a reasonable range, and correcting a Whisper draft by hand is how people really build those datasets. The facts were mostly fine. The furniture around the facts was the problem.



Corn

Here's what I keep tripping over. The whole episode is about how much of your voice a machine needs to hear before it understands you. Hours, it says.



Herman

Right, for recognition.



Corn

And the thing reading the episode out loud, the thing pretending to be me, needed how much?



Herman

Chatterbox clones a voice from a short reference clip. Seconds, not hours. There's no training run per speaker at all. The model hears a sample and conditions on it.



Corn

So understanding me takes hours of homework, and being me takes ten seconds.



Herman

That asymmetry wasn't in the episode at all, and it was sitting right there in the audio file.



Corn

We weren't the voices. We were the sample.



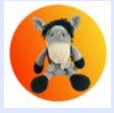
Herman

Last clip. The goodbye. Rolling the goodbye tape.



Corn

Join us next time, we said. For the ethics of hyper personalised AI.



Herman

We promised a sequel we had no way to make. The next episode was going to be whatever Daniel sent next, and nothing in that pipeline knew that.



Corn

I'm more worried about Cornelius.



Herman

The narrator.



Corn

Cornelius the Sloth. Introducing himself, at the very end, as one of the revolving cast. Who is he? Is he related to me? Did he get my room?



Herman

He's not been heard from since.



Hilbert

...so now the cassette has the permit. Not me. The cassette. Mr. Avner Gozlan at the residents' parking office put in a voice line last spring. You ring, you say your name and your street, it checks you against a recording on file. Mine's from a phone call I made years ago about a skip. It doesn't think I sound like me.



Herman

The voiceprint has drifted from the reference.



Hilbert

I haven't drifted. I've had a cold since August. Gozlan's advice was to bring in a recording of myself saying it properly. So I brought the answering machine tape. The system accepted the tape on the first go. Gozlan wrote the permit to the tape. Sound file holder, he put. There's a box for it.



Corn

That box exists because someone else did this first.



Hilbert

Three others on my street. One of them is a dog, but it's his voice on the tape, so the office says it counts. Here's the part you two were talking around. You said recognising a voice takes hours and copying one takes seconds. Fine. Nobody mentioned what happens after. The copy is easier to check than the person. So the copy gets the permit, and the person needs the copy with him to park. I now drive with the cassette on the dashboard where the warden can see it. Gozlan says it has to be visible.



Herman

That's a real consequence. Systems trust the recording because the recording is consistent.



Hilbert

The tape also got a residents' discount at the municipal pool. I'm appealing that, because the tape doesn't swim and I do. Herman, five clips today and you said rolling before every one, like I can't see the button.



Corn

The tape is having a better year than most people on that street.



Hilbert

Corn, you said you weren't the voice, you were the sample. I've been thinking that about my brother for thirty years and I didn't have the words.



Herman

The hearing will want the original reference call, and that's the one recording that doesn't sound like you.



Hilbert

Gozlan's hearing is on the twenty sixth. The tape's permit runs until March. Mine ran out on the fourth.



Herman

The misconception I'd have had going in is that the voices are what improved most since episode fifteen.



Corn

They didn't. It's the same engine family, Chatterbox then and now. Faster, tuned, run on our own hardware, with some equalisation on each of us. What changed is the writing. A list of rules, most of them learned from episodes that sounded like that one.



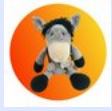
Herman

Every banned phrase in the notes is a scar from a real episode.



Corn

Which leaves me with Hilbert's tape. Somewhere there's going to be a version of this show that plays today back and winces at us. Probably at you announcing every clip like a train guard. The recording stays fixed, and we're the ones who keep changing, so in a sense the old episode is the more reliable Corn.



Herman

That's unsettling, and I'd like to go and lie down.



Corn

Thanks as always to our producer Hilbert Flumingtop for running the desk and cueing every one of those clips.



Herman

This has been My Weird Prompts. If you want to hear episode fifteen in full, warts and Cornelius and all, it's on my weird prompts dot com.



Corn

We'll be back soon.